Reliable activation to novel stimuli predicts higher fluid intelligence

Matthew J. Euler a,b,⁎, Michael P. Weisend c,2, Rex E. Jung c,d, Robert J. Thoma c,e, Ronald A. Yeo b

a Department of Psychology, University of Utah, 380 S. 1530 E. BEHS 502, Salt Lake City, UT 84112, USA
b Department of Psychology, University of New Mexico, MSC03 2220, Albuquerque, NM 87131, USA
c The Mind Research Network, 1101 Yale Blvd. NE, Albuquerque, NM 87106, USA
d Department of Neurosurgery, University of New Mexico, MSC10 5615, Albuquerque, NM 87131, USA
e Department of Psychiatry, University of New Mexico, MSC 09 5030, Albuquerque, NM 87131, USA

A R T I C L E   I N F O

Article history:
Received 17 November 2014
Accepted 28 March 2015
Available online 8 April 2015

Keywords:
Fluid intelligence
Time-frequency analysis
Inter-trial phase clustering
Neural variability

A B S T R A C T

The ability to reliably respond to stimuli could be an important biological determinant of differences in fluid intelligence (Gf). However, most electrophysiological studies of Gf employ event-related potential (ERP) measures that average brain activity over trials, and hence have limited power to quantify neural variability. Time-frequency analyses can capture cross-trial variation in the phase of neural activity, and thus can help address the importance of neural reliability to differences in Gf. This study recruited a community sample of healthy adults and measured inter-trial phase clustering (ITPC), total spectral power, and ERP amplitudes elicited by Repeated and Novel non-target stimuli during two visual oddball tasks. Condition effects, relations among the EEG measures, and relations with Gf were assessed. Early visual responses to Repeated stimuli elicited higher ITPC, yet only ITPC elicited by Novel stimuli was associated with Gf. Analyses of spectral power further highlighted the contribution of phase consistency to the findings. The link between Gf and reliable responding to changing inputs suggests an important role for flexible resource allocation in fluid intellectual skills.

© 2015 Elsevier Inc. All rights reserved.

Introduction

Electrophysiological research on intelligence has demonstrated a number of associations between features of event-related potentials (ERPs) and variation in intellectual skills. Perhaps the most consistent finding is the inverse relation between intellectual ability and the latency of the P300 component elicited by infrequent stimuli during oddball paradigms (Bazana and Stelmack, 2002; Beauchamp and Stelmack, 2006; De Pascalis et al., 2008; Liu et al., 2011; Troche et al., 2009; Walhovd et al., 2005). A number of these and other studies have also shown a positive relationship between P300 amplitude and intelligence (Bazana and Stelmack, 2002; Beauchamp and Stelmack, 2006; De Pascalis et al., 2008; Troche et al., 2009; Wronka et al., 2013), though some studies have not observed this relationship (e.g., Peltz et al., 2011). Similarly, studies examining ERP correlates of inspection time have demonstrated a number of associations between intelligence and early ERP components, with different studies variously identifying component amplitudes, latencies, or wave shape characteristics as being the most predictive (see Deary, 2000, pp. 280–282; and Hill et al., 2011, for reviews).

Despite the overall success of the ERP approach, one limitation of these measures is that they inherently involve averaging neural responses over many trials. Although it is often implicitly assumed that averaged ERPs only reflect activity that is reliable across trials, in practice, effects such as trial-to-trial latency jitter can distort the resulting averages (Luck, 2014, p. 58). Thus, as several authors have pointed out (Barrett and Eysenck, 1992; Batt et al., 1999), ERP measures are necessarily limited in their ability to capture aspects of neural variability that might be of theoretical interest as potential biological mechanisms of intellectual variation. The significance of behavioral intra-individual variability to variation in intelligence is well-established through research on decision time variation (Jensen, 1992) and other measures (Fox et al., 2009; MacDonald et al., 2006; Stuss et al., 2003). At present however, there has been relatively little work investigating how neural variability (or its counterpart — reliability) might underlie variation in intelligence.

Going back over a decade, a number of previous studies investigated relations between neural variability and intelligence using the ERP-based “string length” and associated “variance” measures (Bates et al., 1995; Batt et al., 1999; Robinson, 1997). These indices aimed to assess the presence and significance of “errors” in neural transmission for differences in intelligence (Deary and Caryl, 1997). In essence, they quantified the length of the average evoked potential wave, with the
rationale being that greater variability in neural transmission created greater variability in the waves elicited on individual trials, in turn leading to a smoother, less complex average wave in lower ability individuals, and thus a shorter string length (Hendrickson and Hendrickson, 1980). Although variability in neural responding is conceptually appealing as a variable that might differentiate high and lower ability individuals at the biological level, as recently recounted by Mackintosh (2011), there are a number of methodological concerns related to the string length measure (Robinson, 1997). After several groups found either zero or negative associations between string length and intelligence (Barrett and Eysenck, 1994; Batt et al., 1999), the measure largely fell out of favor. Thus, following the challenges to classic ERP measures of neural variability, the current ERP-intelligence literature is largely characterized by amplitude and latency effects that represent averaged responses over trials, with only a small number of reports investigating other measures (e.g., Jausovec and Jausovec, 2000; Saville et al., 2011).

In recent years significant progress has been made in the field of electrophysiology, particularly in terms of the increasing sophistication and ease of investigating time-frequency characteristics of neural responses (Makeig et al., 2004). These measures provide distinct information from ERPs, largely due to their ability to quantify variability in the phase of neural responses over trials, or the consistency of neural response timing in specific frequency bands and time points relative to experimental events. The importance of phase variation in neural dynamics is increasingly being recognized. Various phase-dependent measures are under investigation as possibly providing fundamental mechanisms for coordinating information processing in the brain (Buzsáki and Draguhn, 2004; Canolty and Knight, 2010; Varela et al., 2001), and abnormal phase dynamics differentiate healthy individuals from patient groups with various neurological and psychiatric conditions (Uhlhaas and Singer, 2006, 2010). In the case of cross-trial phase consistency, the inter-trial phase clustering measure (ITPC) quantifies the consistency of signal phase angle over trials, at each given time-frequency point relative to experimental events. Conceptually then, ITPC provides information about the overall consistency or uniformity of the neural response over trials, and thus may provide a more robust way to operationalize neural reliability.

Based on these considerations, the present study sought to revisit the construct of neural reliability in relation to intelligence, with a focus on the ITPC measure. The rationale for our specific approach is as follows. Given that efficient coordination of early perceptual circuits may be necessary for effective stimulus evaluation and responding, and the substantial literature supporting relations between basic sensory discriminations and intelligence (Jensen, 2006; Melnick et al., 2013), we reasoned that greater reliability in early perceptual processing might be especially associated with intelligence. We focused our analyses on fluid intelligence (Gf) in particular on the expectation that dynamic brain activity would relate most strongly to fluid rather than crystallized skills or broader constructs such as general intelligence (see for example, Neubauer and Fink, 2003). To assess ITPC’s sensitivity to factors that might influence neural reliability, we manipulated stimulus familiarity in an effort to produce varying levels of demand on the early perceptual system. Last, to evaluate the specificity of ITPC to neural reliability apart from related measures and constructs, we conducted analogous tests of our hypotheses using ERPs and time-frequency measures of neural activation, and also assessed the extent of overlap among the measures.

To those ends, we studied early perceptual responses to visual stimuli as quantified by ITPC, total event-related spectral power (“total power”), and ERP amplitudes. Neural activity was elicited by non-target stimuli during a classic, two-stimulus, visual oddball task (involving many Repeated non-targets and infrequent unfamiliar target stimuli) and during an “inverted oddball” task (involving exposure to many Novel non-target stimuli interspersed with infrequent familiar targets). This design enabled examination of neural reliability in response to a single Repeated stimulus versus continuously changing stimuli, and hence a test of the relations between consistent responding and intellectual ability at different levels of demand on the early perceptual system. The analogous comparisons with total power and ERP amplitudes sought to clarify the extent of unique information provided by ITPC.

Based on the expectation that reliable responding would be most facilitated by Repeated stimulus exposure and less facilitated under conditions of continuously changing stimuli (because different stimuli should recruit slightly different networks within the perceptual system; Cox and Savoy, 2003; Tanaka, 1996), we hypothesized that Repeated stimuli would elicit greater ITPC than Novel stimuli. Along the same lines, we hypothesized that while greater ITPC in general may predict higher Gf this should be especially true in the Novel stimulus condition. Because total power also captures non-phase-locked activation, and ERPs can be distorted by latency jitter (which reduces their sensitivity to phase-locking per se), we hypothesized no differences between Repeated and Novel stimuli for those measures, and weaker relationships between those measures and Gf. Finally, in light of the expected overlap between the three measures of stimulus-related activity all acquired from the same data, we conducted a series of planned, follow-up analyses to further assess the uniqueness of any observed effects.

Materials and methods

Participants

Study participants were recruited from the local community through an existing database of individuals who had participated in prior studies of intelligence and creativity (Jung et al., 2010a,b), and from an advertisement posted on an internet classified website. All study procedures were reviewed and approved by the University of New Mexico Health Sciences Center Human Research Protections Office. Participants were screened for neurological or psychiatric illnesses, and to ensure they had normal or corrected to normal vision. Participants were compensated $15 per hour for their time. Thirty individuals between the ages of 18 to 30 participated in the study. Participants had between 11 and 18 years of formal education (μ = 14, SD = 2.2), and were on average 23 years old at the time of cognitive testing and EEG recording. Following EEG data processing described below, a total of 29 participants (10 women) had sufficient numbers of artifact-free trials in both conditions, and were included in the final study dataset.

Stimuli

Stimuli consisted of 411 black and white line drawings depicting common objects (animals, tools, musical instruments, vehicles, etc.) which were taken from the larger set of 520 stimuli developed by the Center for Research in Language’s International Picture Naming Project (Bates et al., 2003). Both the full set and subsets of these stimuli have been widely utilized in fMRI research (e.g., Saccuman et al., 2006; Stark and Squire, 2000) and EEG research on repetition priming (Gruber and Müller, 2005). The specific subset selected for the study was chosen on the basis of comparable visual complexity as assessed by digital file size, and comparable naming reaction times (Szekely et al., 2004). For each participant, a pseudo-random number generator was used to select a unique set of 289 stimuli from the larger subset.

Experimental session

Participants were seated in a sound and light-attenuated room for EEG acquisition. The stimuli consisted of black drawings on a white background and were in turn centrally presented on a larger gray background on a monitor located approximately 1.5 m in front of the participant. Drawings subtended 3.2° by 3.0° of visual angle. Participants completed a four-minute eyes-closed resting period prior to the
experimental tasks, each of which lasted approximately 11 min. The order of experimental tasks was counterbalanced across participants.

**Classic and inverted visual oddball tasks**

Study tasks consisted of classic and “inverted” visual oddball paradigms, respectively. In each task, participants were shown 240 non-target stimuli that were randomly interspersed with 48 target stimuli, to which they were instructed to respond as quickly as possible. Participants responded via a button press with their right index finger, using a custom-made, glove-shaped, fiber-optic response device. For clarity, we will refer to the target stimuli in both conditions as “Unexpected” to avoid confusion with the perceptual stimulus dimensions of Repeated vs. Novel drawings. In the classic oddball task, a single specific stimulus was presented on all non-target trials (Repeated non-targets) and was randomly interspersed with 48 unique target stimuli selected at random from the larger subset (Unexpected Novels). During the inverted oddball task, 240 unique Novel stimuli were presented (Novel non-targets), which were randomly interspersed with 48 trials in which the Repeated stimulus was shown (Unexpected Repeats). Due to counterbalancing, all participants were briefly shown their particular Repeated stimulus prior to beginning either task. Participants were instructed to respond to Unexpected Novels in the classic oddball task and Unexpected Repeats in the inverted oddball task. For each participant, the Repeated stimulus was chosen at random and held constant across the two tasks. The order of stimulus presentation for all trials in each task involved presentation of a white fixation cross which was randomly jittered 1000–2500 milliseconds, followed by a standard or target stimulus for 1000 milliseconds. In target conditions, the participant’s response ended the trial. Participants were instructed to maintain fixation on the cross in order to minimize eye movements, and were given an opportunity to rest every 50 trials.

**Cognitive testing**

Following EEG data collection, most participants without prior testing (n = 15) were administered select subtests of the Wechsler Adult Intelligence Scale III. All study participants ultimately completed the Picture Completion, Vocabulary, Digit-Symbol Coding, Similarities, Block Design, Arithmetic, Matrix Reasoning, Digit Span, Information, Picture Arrangement, Symbol Search, and Letter-Number Sequencing subtests (Wechsler, 1997). Due to an error all but one participant also completed the Object Assembly subtest; this single missing score was replaced using the expectation–maximization procedure available in SPSS.

**Electrophysiological recordings**

Individual EEG data was obtained using the 128-Channel Biosemi Active-Two amplifier System (Van Rijn et al., 1990). Data were sampled on-line at 1024 Hz with a 0.16–100 Hz band-pass filter. Vertical and horizontal electrooculograms (VEOG, HEOG) as well as EKG were monitored via six additional channels placed above and below the left eye, at the external canthi of each eye, and near the clavicles. An additional electrode placed on the nose-tip was used as the offline reference.

**EEG pre-processing, artifact correction, and removal**

EEG data were pre-processed using the ASA-lab software package (version 4.7.12, ANT Neuro, Enschede, The Netherlands), as well as custom routines running in MATLAB (R2012b, Natick, MA). Scalp maps were created using the EEGLAB topoplot function (Delorme and Makeig, 2004). Continuous EEG data were first loaded into ASA-lab, re-referenced to the nose reference, and down-sampled to 512 Hz to speed further processing. A high-pass filter at 0.3 Hz was then applied to correct for any remaining low-frequency artifacts, which is within or below the range adopted by previous studies in this area (ranging from 1.0–3.0 Hz; Batt et al., 1999; Beauchamp and Stelmack, 2006; Hill et al., 2011; Jaušovec and Jaušovec, 2000; Troche et al., 2009). Non-target trials were extracted and divided into stimulus-locked epochs that spanned −1000 to 1000 milliseconds pre to post-stimulus. Artifact removal was accomplished using several methods. First, prior to high-pass filtering, the data was corrected for ocular artifacts using the topographical PCA-based method implemented in ASA-lab (Ille et al., 2002). Individual epochs were then visually inspected for artifacts due to facial and other muscular activity, or other sources of noise, and manually excluded prior to being exported for further processing. Finally, prior to time-frequency analysis, custom-built MATLAB routines removed any remaining epochs with amplitudes exceeding an absolute threshold of ±75 mV, or with voltage changes exceeding ±60 mV within 100 milliseconds windows (Luck, 2005).

**ERP analyses and grand averaged scalp topographies**

Epochs that survived the artifact-correction procedures were retained for ERP and time-frequency analyses that were conducted using custom routines running in MATLAB 8.0. Since ITPC values are influenced by the number of trials included in the final calculation (i.e., too few trials can erroneously inflate it; Cohen, 2014), we chose 120 trials as the final number to retain for all subjects in each condition, which balanced the need to have representative measures of neural responses while also maximizing the final sample size. For this same reason, stimulus-related activity from the two target conditions was not analyzed in this study, since fewer than 25 trials would be available across all conditions and participants (after removing those contaminated by artifacts), thus impeding reliable estimates of those ITPC measures, or valid comparisons to the non-target conditions. For ERP analyses, epochs were low-pass filtered at 30 Hz, and baseline-corrected using the mean amplitude between −300 to −100 milliseconds pre-stimulus. Given the relatively limited spatial resolution of EEG, we chose a data-driven approach to quantifying stimulus-related activity and extracted the study measures from the sensors where the early perceptual response was maximal in the ERP. For that purpose, the filtered, baseline-corrected data were averaged across epochs, participants, and conditions, to create grand-averaged topographic maps of the neural response to the non-target stimuli. Visual inspection indicated that the early response over occipito-parietal sensors peaked at approximately 120 milliseconds (corresponding to the P1 component; Luck, 2014) and was maximal at or approximating PO8, PO9, and PO10 in the extended International 10–20 system (see Fig. 1; Supplementary Fig. 1 depicts grand-averaged ERPs from all four task conditions). Averaged values across those four sensors were in turn used for all ERP and time-frequency analyses. For ERP analyses, epochs were averaged over trials and then sensors, separately by condition. Based on the peak latency observed in the grand average, mean amplitudes were calculated for each condition and subject within a window from 95–145 milliseconds post-stimulus.

**Maximum ITPC and time-frequency power**

The unfiltered, non-baseline-corrected epochs processed in the ERP analyses above were retained for time-frequency analyses. Except for the specific study parameters and where otherwise noted, computation of all time-frequency variables proceeded as detailed in the recent text by Cohen (2014). In the first step, the data were averaged over the four sensors of interest, yielding an average matrix of trials by samples.
for each participant and condition. The complex representations of the EEG signals were then obtained via the S-transform (Stockwell et al., 1996) with a resolution of 200 frequency bins of 0.5 Hz width from 0.5–100 Hz over all 1024 samples (2000 milliseconds) in each epoch. For ITPC, the instantaneous phase angles of the complex signal in each trial were extracted and ITPC was then calculated according the formula:

\[
ITPC = \left| n^{-1} \sum_{t=1}^{n} e^{i \phi(t,f)} \right|
\]

where \( ITPC \) equals the length of the average vector of the complex polar representations of phase angles \( e^{i \phi} \) at each time-frequency point and trial \( f,t \) across all \( n \) trials as detailed in Cohen (2014). This measure quantifies the extent of consistency or “clustering” within the cross-trial distribution of phase angles at each given time-frequency point, and assumes a value between 0 and 1. A value of 0 indicates a fully uniform distribution of phase angles over trials (at that given time-frequency point), while a value of 1 indicates fully identical phase angles (Cohen, 2014).

The identical single-trial complex representations from the S-transform were used to compute total power, which was calculated by multiplying the complex result at each time-frequency point and trial \( f,t \) across all \( n \) trials as detailed in Cohen (2014; Cohen et al., 2008). This measure
developed via the S-transform (Stockwell et al., 1996) with a resolution of 200 frequency bins of 0.5 Hz width from 0.5–100 Hz over all 1024 samples (2000 milliseconds) in each epoch. For ITPC, the instantaneous phase angles of the complex signal in each trial were extracted and ITPC was then calculated according the formula:

\[
ITPC = \left| n^{-1} \sum_{t=1}^{n} e^{i \phi(t,f)} \right|
\]

where \( ITPC \) equals the length of the average vector of the complex polar representations of phase angles \( e^{i \phi} \) at each time-frequency point and trial \( f,t \) across all \( n \) trials as detailed in Cohen (2014). This measure quantifies the extent of consistency or “clustering” within the cross-trial distribution of phase angles at each given time-frequency point, and assumes a value between 0 and 1. A value of 0 indicates a fully uniform distribution of phase angles over trials (at that given time-frequency point), while a value of 1 indicates fully identical phase angles (Cohen, 2014).

The identical single-trial complex representations from the S-transform were used to compute total power, which was calculated by multiplying the complex result at each time-frequency point and trial \( f,t \) across all \( n \) trials as detailed in Cohen (2014; Cohen et al., 2008). The baseline period was defined from \(-500–200 \) milliseconds pre-stimulus to avoid including any stimulus-related activity that smeared into the baseline due to the time-frequency transformation. Visual inspection of the grand averaged plots primarily indicated event-related activity in the early post-stimulus period in the classical theta (4–8 Hz) and (alpha 8–12 Hz) frequency bands (see Fig. 2). For use in follow-up analyses (described below), phase-locked “evoked” power was also calculated. This involved subtracting the time-domain ERP from each single trial, obtaining the complex representations and multiplying by the conjugate, averaging over trials, and baseline-correcting as above to obtain non-phase-locked power. The resulting value was then subtracted from baseline-corrected total power to obtain phase-locked power.

Maximum values for ITPC and power were used to operationalize the dependent variables since the choice of a maximum value within a window allows for individual differences in the peak frequency and latency of the response, while still quantifying the peak reliability and magnitude. Based on the grand-averaged plots, the maximum values of ITPC and total power within a window from 4–12 Hz and 70–250 milliseconds post-stimulus were retained for time-frequency analysis of the early perceptual response.

### Statistical analysis

Analyses of condition effects on ERP and time-frequency variables were conducted in SPSS version 20. To anticipate any concerns regarding effects of potential outliers on tests of associations with Gf, correlation analyses of the primary hypotheses regarding ITPC, total power, and ERP amplitudes were performed in MATLAB using the robust correlation toolbox, set to the default values (Pernet et al., 2013). In particular, we conducted the percentage-bend correlation as well as the skipped correlation (Rousseeuw, 1984; Rousseeuw and Driessen, 1999; Verbomen and Hubert, 2005) between Gf and all six EEG variables of primary interest (Maximum ITPC, Maximum Power, and Mean ERP Amplitudes for the Repeated and Novel conditions). The percentage-bend and skipped correlations respectively control for the effects of univariate and bivariate outliers on the estimated associations (Wilcox, 2005). The toolbox automatically implements the Bonferroni adjustment for multiple comparisons for each test (i.e., \( \alpha = .05/6 = .008 \)) and provides bootstrapped confidence intervals for the correlations themselves, thereby providing stringent tests of the main hypotheses. Because the goal of the follow-up analyses was to provide broader information regarding the relations between the variables under study, for those analyses we report conventional Pearson correlations, unadjusted for multiple comparisons.

### Results

#### Principal analyses

**Fluid intelligence composite**

In addition to Gf, WAIS-III FSIQ scores were calculated to facilitate comparisons with the broader literature. The average FSIQ score for the study sample was 112 (SD = 15.46, range = 83–138). The sample was not skewed with respect to FSIQ scores (Shapiro–Wilk Statistic = .964, \( p = .412 \)); and four participants scored within a point of the low average range or lower and four participants performed in the very superior range. Male and female participants did not differ with respect to FSIQ (\( F(1,28) = .446, p = .51 \)). Raw scores from all but the Vocabulary, Similarities and Information subtests were converted to z-scores and then summed to provide the composite measure of Gf. The internal consistency reliability of the composite was very good (Cronbach’s \( \alpha = .819 \)), and demonstrated a smaller correlation with crystallized skills as measured by the WAIS Verbal Comprehension
Index ($r(29) = .705, p < .001$) than did FSIQ ($r(29) = .889, p < .001$), indicating that it provided a robust index of Gf.

**Task performance**

Participants were highly accurate in detecting target stimuli in both conditions (Unexpected Novels: $\mu = 98.7\%$ correct; Unexpected Repeats: $\mu = 98.8\%$ correct), with no difference in accuracy across conditions ($F(1,28) = .015, p = .904$). Average reaction time was significantly faster in the Unexpected Repeated condition ($\mu = 480.96$ ms, SD = 74.63) relative to the Unexpected Novel condition ($\mu = 503.03$ ms, SD = 75.83; $F(1,28) = 4.318, p = .047$), indicating that participants were faster to respond to Unexpected Repeated targets presented among Novel non-targets, relative to the converse. The correlation between Gf and average reaction time in the Unexpected Novel ($r(29) = -.32, p = .09$) and Unexpected Repeated condition ($r(29) = -.21, p = .28$) were in the expected negative direction but were not statistically significant.

**Condition effects on ITPC, total power, and ERPs**

Descriptive and inferential statistics for ERP amplitudes and time-frequency values are listed in Table 1. Consistent with the hypothesis that Repeated stimuli would elicit more reliable responding over trials than changing stimuli, a one-way repeated measures ANOVA indicated higher maximum ITPC in the Repeated relative to the Novel condition ($F(1,28) = 7.176, p = .012$, partial $\eta^2 = .204$). No differences were observed between conditions for maximum power ($F(1,28) = 1.126, p = .298$) or ERP amplitudes ($F(1,28) = .114, p = .738$), suggesting that ITPC better captured the hypothesized condition effect than overall neural activation or ERPs. Follow-up analyses of the maximum ITPC effect indicated that ITPC peaked somewhat later ($F(1,28) = 6.852, p = .014$, partial $\eta^2 = .197$) and at a non-significantly higher frequency ($F(1,28) = 3.043, p = .092$) in the Repeated compared to the Novel condition. Conditions did not differ with respect to the maximum latency or frequency of total power.

**Stimulus-related activity and Gf**

As shown in Table 2, robust correlation analyses indicated that only the association between ITPC in the Novel condition and Gf was significant when controlling for multiple comparisons. These results were unchanged when each correlation was tested alone (i.e., uncorrected), although in that instance, the percentage-bend correlation ($r_{pb}$) between Gf and ERP amplitudes in both conditions approached significance, as did the percentage-bend correlation for Gf and Novel Power. When examined as single tests, Novel ITPC consistently accounted for 29% or more of the variance in Gf (Pearson Skipped Correlation $r(29) = .54$; $r_{pb}(29) = .56, p = .0016$; and see scatter plot in Fig. 3).

**Follow-up analyses**

**Specificity of ITPC**

In order to better understand the extent to which ITPC captured unique features of neural responses apart from the other measures of stimulus-related activity, several planned follow-up analyses were...
Conducted. To assess the extent of shared variance among the six EEG variables in the study, the conventional Pearson correlations are presented in Table 3.

As expected, ITPC and total power elicited within the same task were generally the most highly correlated variables, followed by the correlations between variables of the same type across tasks. In that latter instance, the ERP measures were the most highly correlated, while the ITPC values showed the weakest association across conditions. Supplementary analyses of theta and alpha band-specific activity conformed to the overall pattern of results.

**Similarity of ITPC to phase-locked power**

Isolating only the phase-locked or “evoked” power that is subsumed within total power (which captures both phase-locked and non-phase-locked stimulus related activity over trials), is another common method of capturing phase-locked activity in time-frequency representations. Because both ITPC and evoked power emphasize cross-trial phase-locking in their computation, they are expected to be highly correlated (Cohen, 2014; p. 261). Due to this redundancy ITPC was selected to operationalize neural reliability over evoked power, because it quantifies the cross-trial consistency of response timing directly, rather than via the amplitude of the phase-locked activity. Nonetheless, because the two measures both emphasize consistent activation they should demonstrate the same pattern of effects whenever phase consistency is the operative variable driving the results. Consistent with these expectations, the two measures correlated essentially at unity for both tendency is the operative variable driving the results. Consistent with these

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Robust correlations with fluid intelligence.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel ITPC</td>
<td>.594</td>
</tr>
<tr>
<td>Novel power</td>
<td>.338</td>
</tr>
<tr>
<td>Novel ERP</td>
<td>.377</td>
</tr>
<tr>
<td>Repeated ITPC</td>
<td>.199</td>
</tr>
<tr>
<td>Repeated power</td>
<td>.191</td>
</tr>
<tr>
<td>Repeated ERP</td>
<td>.377</td>
</tr>
</tbody>
</table>

| Note: CI = Bonferroni adjusted, 95% Bootstrapped Confidence Interval; Boldface indicates CI does not include zero. |
individuals during low-demand tasks (based on within person difficulty level; Dunst et al., 2014), but relatively higher activation during more difficult tasks (Neubauer and Fink, 2009). Given that the relation between neural reliability and Gf only held for Novel stimuli (despite higher ITPC in the Repeated condition), the present results appear to be generally consistent with the NE hypothesis. Specifically, they conform to the pattern predicted by NE—that a positive relation between neural activity and cognitive ability should be preferentially observed in more demanding tasks. Coordinating neural assemblies is presumably metabolically costly, and maintaining high reliability even in low-demand conditions may not be especially adaptive. Conversely, because rapid registration of novel percepts might be associated with evolutionarily significant events (e.g., threats) the ability to respond reliably in those instances may exemplify the adaptive expenditure of neural resources.

A more specific consideration regarding NE relates to the meaning of signal phase and the ITPC measure. As discussed by Cohen (2014) phase refers to the timing of neural responses (in a particular frequency band) relative to experimental events and can be thought of as reflecting a “functional configuration” of the underlying population (p. 268). In turn, since ITPC quantifies the consistency with which a given configuration is achieved over trials, the effect of higher ITPC in the Repeated condition suggests that, consistent with hypotheses, the early visual system is more readily configured when percepts are unchanged. However, because Repeated stimulation involves lower perceptual demands, this situation may not elicit especially meaningful differences in the more adaptive capacity to functionally configure visual networks in general. Ultimately, given that the conditions only differed with regard to phase consistency but not overall activation, it is unclear how closely ITPC tracks neural resource expenditure, and thus the meaning of efficiency with regard to reliability measures. (See Supplementary material for further results and discussion relevant to the NE hypothesis).

### Methodological implications

More broadly, the overall pattern of results raises several considerations regarding the meaning of neural reliability and activation, and of the inter-relationships among various EEG measures of neural activity. To briefly reiterate, ITPC and total power exhibited a similar pattern of associations with Gf across conditions, with relations for power being weaker and non-significant. Neither ITPC nor total power in the Novel stimulus condition was especially highly correlated with its counterpart measure in the Repeated condition. On the other hand, ERP amplitudes were highly correlated across conditions and exhibited weak but consistent trends toward associations with Gf. The stronger effects for ITPC relative to total power are consistent with a recent computational modeling study which showed that ITPC is more likely to detect changes in phase-locked stimulus-related activity than is total power, simply as a function of its statistical properties (Ding and Simon, 2013). Crucially however, that study also demonstrated that evoked power exhibits nearly identical statistical properties to ITPC, and performed as well or better as a function of signal-to-noise ratio. Thus, those results together with our own highlight the critical contribution of cross-trial consistency in activation rather than activation itself to the present effects, and hence the utility of the neural reliability construct. Given that ERPs are commonly considered to reflect only phase-locked activity in the time-domain (albeit with the caveat noted above), the observed trend-level effects for ERPs may also highlight the greater sensitivity of ITPC to neural reliability per se. Overall, the present results highlight how different analysis approaches emphasize different features of the underlying event-related responses. Here, isolating phase dynamics made for a more sensitive measure of the brain’s flexibility and responsiveness to environmental stimuli among higher Gf individuals.

### Limitations

Some limitations of the present study should also be noted. In particular, although participants exhibited a high degree of accuracy in both tasks, reaction times were faster in response to Unexpected Repeated relative to Novel targets, indicating that the tasks may not have been ideally matched on characteristics that influenced behavioral responses. Despite the relatively repetitive nature of monitoring for a Novel target stimulus amidst Repeated stimuli, there were no differences between conditions in ongoing alpha power (see Supplementary section: Baseline Alpha Power), suggesting that differences in arousal or attentiveness (Foucher et al., 2004; Oken et al., 2006) are unlikely to account for the effects. A second consideration concerns whether higher ability participants might be attending differently or adopting a unique strategy to performing the task at hand (e.g., Deary, 2000, p. 279), and thus whether neural mechanisms related to higher-order processes might actually account for relations with Gf. In the present study, it is possible that the Novel stimulus condition placed greater demands on visual discrimination or working memory resources due to the much greater diversity of non-targets in that condition, which could manifest as top-down modulation of parieto-occipital activity. The fact that there were no differences in ERP amplitudes between the conditions suggests that they were adequately matched in terms of their demands on visual discrimination processes (Vogel and Luck, 2000), and somewhat mitigates this concern. However, since the present study was primarily concerned with perceptual effects, task demands were not explicitly manipulated, making it difficult to fully address the potential role of working memory or other higher-order processes. Future studies should address the

### Table 3

Correlations between ERPs, total power, and ITPC.

<table>
<thead>
<tr>
<th></th>
<th>Novel ITPC</th>
<th>Novel power</th>
<th>Novel ERP</th>
<th>Repeated ITPC</th>
<th>Repeated power</th>
</tr>
</thead>
<tbody>
<tr>
<td>Novel power</td>
<td>.812**</td>
<td>–</td>
<td>.123</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Novel ERP</td>
<td>.302</td>
<td>.123</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Repeated ITPC</td>
<td>.580**</td>
<td>.568**</td>
<td>.014</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Repeated power</td>
<td>.560**</td>
<td>.654**</td>
<td>.027</td>
<td>.883**</td>
<td>.174</td>
</tr>
</tbody>
</table>

**p < .003; * = p < .05; all others p > .10.**

---

Fig. 3. ITPC elicited in response to Novel stimuli and Gf.
possible impact of task demands and higher-order processes on the relation between neural reliability and Gf.

Finally, given the relatively low spatial resolution of scalp EEG, it is difficult to directly relate the present results to the Parieto-Frontal Integration Theory of intelligence (P-FIT; Jung and Haier, 2007), which is the other major theory in this literature. Nonetheless, P-FIT components in extrastriate and fusiform occipital areas may well be implicated given their role in visual attentional and sensory discrimination processes during the time course of the P1 and N1 ERP components (Luck, 2014, pp. 75–77; Pratt, 2011, p. 107), which would also be consistent with a recent fMRI study of P-FIT connectivity that demonstrated a critical role of occipital structures in cognitive networks “responsible for the decision making process during fluid reasoning” ( Valkhvit et al., 2014).

Ultimately however, future studies with higher spatial resolution will be needed to identify the generators of the current effects and their connection to current neuroanatomical theories.

Conclusions

In summary, the present results support the use of ITPC as a measure of neural reliability that provides greater sensitivity to the consistency of neural dynamics relative to ERPs or time-frequency measures of overall activation (see Saville et al., 2011; for a contemporary ERP-based measure of neural variability). Higher neural reliability to Novel stimuli was uniquely associated with Gf, suggesting that higher ability individuals exhibit greater flexibility in their capacity to efficiently coordinate visuoperceptual networks. Future work in this area should clarify the impact of various perceptual vs. cognitive demands on neural reliability, and its broader connection to efficiency and neuroanatomical models of intelligence.

Acknowledgments

The authors would like to thank Ranee Flores and Eva Padilla for assistance with participant recruitment and data collection.

Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.neuroimage.2015.03.078.

References


